CRÍTICA, *Revista Hispanoamericana de Filosofía*
Vol. XXIII, No. 67 (abril 1991): 41–58

# TEORÍA GENERAL DE LAS DECISIONES

ADOLFO GARCÍA DE LA SIENRA
Instituto de Investigaciones Filosóficas
Universidad Nacional Autónoma de México

Mi propósito en el presente trabajo no es proporcionar una reconstrucción sistemática de lo que normalmente se conoce como teoría de las decisiones, sino introducir un esquema más amplio, de aplicabilidad más general, que contiene como caso especial a esta teoría. En la vida humana (en particular en el manejo de una empresa en una economía de mercado) encontramos frecuentemente situaciones de decisión (con o sin incertidumbre) en las que es necesario entrar en un proceso más o menos complicado de evaluación antes de adoptar una de las posibles opciones que se presentan. La economía neoclásica dio surgimiento a un concepto de proceso de decisión en el que se tiene un espacio matemáticamente representable de decisiones, junto con la posibilidad de maximizar dentro de este espacio una meta particular, como el beneficio o la utilidad; *i.e.* es posible *computar* dentro de este espacio aquel o aquellos puntos que maximizan el objetivo buscado. En estos procesos de decisión se considera que el agente que toma las decisiones es racional syss, sabiendo cuales son las decisiones óptimas, adopta esas y no otras. Simon (1976) ha llamado a esta forma de adoptar decisiones 'racionalidad sustantiva'. Este mismo autor ha opuesto a la racionalidad sustantiva otra forma de tomar decisiones, que el llama 'comportamiento procesalmente racional', la cual consiste en tomar decisiones como resultado de una “deliberación apropiada”. La primera pregunta que quiero considerar aquí es la siguiente: ¿Es posible caracterizar en términos generales la racionalidad procesal? Y, si esto fuere así, ¿se reduce la racionalidad sustantiva a ser un mero opuesto de ella? Trataré de mostrar, mediante el concepto de estructura general de decisiones, que si no es posible caracterizar de manera exhaustiva y plenamente general la racionalidad procesal, por lo menos es posible introducir un esquema dentro del cual es posible plantearse problemas de racionalidad procesal, e incluso postular algunas restricciones para los mismos. No es difícil ver que la racionalidad sustantiva es un caso particular de este esquema, a saber, el caso en el que es posible *medir* determinados parámetros teóricos y adoptar como reglas de decisión ciertos principios clásicos. No habré de proporcionar aquí, sin embargo, los detalles que muestran cómo se obtiene la teoría especial a partir de la más general.

## 1. *Planteamiento del problema*

Una forma de abordar nuestro problema consiste en proceder a dar una descripción completamente general de las situaciones de decisión. Imaginemos un agente en una situación en que está obligado a tomar una decisión $d$ entre un conjunto fijo y exhaustivo de decisiones $D$ aunque es posible que el agente no tenga conciencia de que es posible adoptar alguna decisión en $D$. Es posible que cada decisión posible $d \in D$ acarree de modo ineluctable “una” consecuencia (una serie de consecuencias) que podemos representar simbólicamente mediante la letra ‘$c$’. En este caso decimos que la toma de decisiones es *determinista* o *sin incertidumbre*. Nuevamente, el agente puede o no saber, para la decisión arbitraria $d$, si la conexión entre $d$ y su consecuencia es determinista, o incluso puede ignorar si hay o no hay tal conexión. Para tener en cuenta este problema de ignorancia, introduciremos una idea muy psicológica de verosimilitud: una escala cualitativa, que puede ser interpretada como grado de creencia, indicando la intensidad de adhesión del agente a la creencia de que una determinada decisión $d$ acarreará una determinada consecuencia $c$. Para dar cuenta del otro problema de ignorancia, diremos que el conjunto $D$ de opciones abiertas al agente no es sino el conjunto de todas aquellas acciones realmente posibles para el agente que él reconoce como tales.

Reinterpretamos así nuestros símbolos como sigue. $D$ es el conjunto de todas las decisiones que el agente reconoce como factibles en la situación. Para cada decisión $d \in D$, el agente reconoce que puede tener lugar una de una serie de posibles consecuencias $c \in C$, y esto de acuerdo con cierto grado de verosimilitud cualitativa. Una vez tomada la decisión $d \in D$ la ocurrencia de la consecuencia $c \in C$ puede ser determinista como se indicó arriba, o también puede depender de ciertas eventualidades. Por ejemplo, el agente (una empresa) tiene que decidir entre fijar uno de dos precios para su producto: alto ($\alpha$) o bajo ($\beta$). La competencia también tiene que fijar uno de estos dos precios, pero el agente no puede saber la decisión que habrá de tomar su competidor en el mercado. Hay dos eventualidades: la competencia fija el precio alto ($A$), o fija el bajo ($B$). Pueden así ocurrir cuatro situaciones: (1) el agente toma la decisión $\alpha$ y la competencia fija el precio alto $(\alpha, A)$; (2) toma la decisión $\alpha$ pero sucede $B$ $(\alpha, B)$; (3) toma la decisión $\beta$ y sucede $A$ $(\beta, A)$; (4) toma la decisión $\beta$ y sucede $B$ $(\beta, B)$. Asociada a cada una de estas situaciones hay una consecuencia. Digamos, en aras del ejemplo, que $(\alpha, A)$ acarrea que el agente gane 150 millones de pesos libres; $(\alpha, B)$ provoca que pierda 50 millones; $(\beta, A)$ hace que gane 100 millones; y $(\beta, B)$ significa que gane 50 millones.

En el evento anterior se ve claramente que la consecuencia resultante de cualquiera de las decisiones adoptadas por el agente es un resultado, final en algún sentido, cuya deseabilidad no depende de la eventualidad que lo produjo. Por ejemplo, es más deseable en cualquier caso tener 150 millones (lo cual ocurre en la eventualidad $(\alpha, A)$) que tener 100 (lo que sucede bajo la eventualidad $(\beta, A)$). Expresaremos en general este hecho diciendo que la relación fundamental de preferencia (o deseabilidad) entre las consecuencias es intrínseca al conjunto $C$ de las mismas. Sin embargo, esto no quiere decir que las cuatro situaciones $(\alpha, A)$, $(\alpha, B)$, $(\beta, A)$ y $(\beta, B)$ sean igualmente preferibles. Ciertamente, algún agente puede estimar que es demasiado peligroso adoptar el precio alto, puesto que el prospecto de ganar 150 millones se ve desmerecido por la posibilidad de perder 50 millones, mientras que la adopción del precio bajo garantiza por lo menos 50 millones de ganancia. Lo que esto significa es que tenemos, además de la relación fundamental de preferencia entre consecuencias, *otra relación de preferencia*, la cual está definida entre las opciones mismas, *i.e.* entre los elementos del conjunto $D$. Por ejemplo, algún agente conservador en cuanto a los riesgos seguramente diría que la opción $\beta$ es preferible a la opción $\alpha$. Determinar la conexión nomológica más general entre la relación fundamental de preferencia y la que tiene lugar entre las opciones es el problema fundamental de la filosofía de la teoría general de las decisiones.

Lo que actualmente se conoce como teoría de las decisiones postula como ley básica que la conexión entre la relación fundamental de preferencia y la que tiene lugar entre las decisiones es la de la utilidad esperada. En otras palabras, sea $u$ la utilidad correspondiente a la relación fundamental de preferencia, y sea $U$ la correspondiente a la otra. La ley de la utilidad esperada afirma que la utilidad $U$ es una mezcla de la utilidad $u$ con las probabilidades de los eventos posibles. En el caso del ejemplo, esto significa que si $p$ es la probabilidad de que la competencia imponga el precio bajo a su producto, entonces $U(a) = 150(1 - p) - 50p = 150 - 200p$ mientras que $U(\beta) = 50(1 - p) + 100p = 50 + 50p$. Se ve que conforme $p$ tiende a cero la utilidad esperada de $\alpha$ crece y tiende a hacerse mayor que la de $\beta$. Cuando la probabilidad de $B$ es .4, la utilidad esperada de $\alpha$ es igual a la de $\beta$, pero uno se pregunta si de verdad alguien que estime en .4 la probabilidad de perder 50 millones está dispuesto a considerar dichas opciones como indiferentes. En otras palabras, alguien podría decir que aunque la utilidad esperada de $\alpha$ sea igual a la de $\beta$, sin embargo, preferiría tomar la decision de adoptar el precio $\beta$ por ser menos riesgosa. En otros términos esto significa que no necesariamente la utilidad esperada es una buena medición de la preferencia sobre las opciones.

El efecto anteriormente descrito ha sido llamado por Watkins (1970) 'el efecto del apostador'. La austeridad y la prudencia burguesas, por ejemplo, recomendarían jamás poner el precio alto y aceptar mejor la posibilidad de obtener al menos 50 millones, en vez de exponerse a perder la misma suma. Esto demuestra que la teoría general de las decisiones no puede adoptar el concepto de utilidad esperada como un concepto fundamental, sino que debe de conformarse, precisamente, con el de preferencia entre opciones.

## 2. *Estructuras de preferencia*

Desde un punto de vista lógico, una relación de preferencia sobre un conjunto $C$ es un subconjunto de $C \times C$, representado mediante el símbolo $\succsim$. Si $c_1, c_2 \in C$, se escribe $c_1 \succsim c_2$ para expresar la proposición de que $c_1$ es al menos tan preferible como $c_2$. Usualmente se supone que la estructura $\langle C, \succsim \rangle$ es un orden débil, *i.e.* que la relación $\succsim$ es conectada, reflexiva y transitiva en C, pero incluso la primera y la tercera de estas suposiciones, las cuales son aparentemente innocuas, pueden restar generalidad a la teoría, pues pueden dejar fuera relaciones concretas, entre consecuencias de situaciones decisorias, que naturalmente deben ser concebidas como relaciones de preferencia de los agentes decisores. Pero estas suposiciones son necesarias para cualquier representación numérica de la relación de preferencia, por lo que el precio que hay que pagar por prescindir de la conectividad o de la transitividad es el el de olvidarse de cualquier función de utilidad para representar la relación de preferencia. La conectividad expresa que el agente siempre sabe, para cada par de consecuencias, si prefiere estrictamente una a la otra o le son indiferentes. La transitividad expresa que la preferencia entre opciones está linealmente ordenada, *i.e.* formando una jerarquía de opciones.

La pregunta relevante aquí es si la conectividad y la transitividad significan una limitación a la generalidad teórica buscada. ¿Es posible que haya una situación decisoria en la que el agente no prefiera una de dos opciones, ni tampoco le sean indiferentes? ¿Es posible que prefiera $c_1$ a $c_2$, $c_2$ a $c_3$, y sin embargo que prefiera $c_3$ a $c_1$? Está claro que experimentalmente es posible establecer —por lo menos— que en ocasiones la transitividad no se cumple por los agentes. Sin embargo, es posible mostrar que en las ocasiones en que tenemos estos círculos no es posible darle significado alguno a la frase 'buena decisión'. Pues en estos casos no hay criterio alguno para decidir cuál es una buena decisión. Supóngase, por ejemplo, que decidimos que la decisión conducente a $c_1$ es preferible a la conducente a $c_2$. En este caso parece razonable escoger $c_1$; pero si se ve con más cuidado, está claro que ello depende únicamente del hecho de que empezamos nuestra consideración por la relación entre $c_1$ y $c_2$. Si hubiésemos empezado por comparar $c_2$ con $c_3$, hubiéramos considerado que es mejor escoger la decisión conducente a $c_2$. Lo que es peor, si hubiésemos empezado comparando $c_3$ con $c_1$ hubiéramos llegado a la conclusión de que lo razonable es escoger $c_3$. En otras palabras, si la transitividad no se cumple entonces cualquier decisión puede ser considerada como "buena". Desde luego, en este caso es imposible argumentar en modo alguno en favor de una u otra decisión. Es por ello que considero que la mínima condición que se debe imponer a la relación de preferencia es la transitividad.

Pero tampoco es necesario presumir que se tiene que cumplir siempre la conectividad: nada impide considerar a la preferencia como un orden parcial sobre el conjunto $C$. En tal caso (suponiendo que se cumple la transitividad), el conjunto de las clases de indiferencia no sería una cadena, sino un retículo, y no habría necesariamente una clase de indiferencia óptima, sino posiblemente varias. Sin embargo, todavía tendría sentido hablar de las "mejores" consecuencias u opciones, pues éstas serían precisamente los elementos de las clases maximales del retículo, suponiendo que toda cadena tuviese un elemento maximal. Así, podríamos caracterizar nuestra estructura generalizada de preferencia simplemente como un orden parcial tal que toda cadena tiene elementos tanto maximales como minimales (esto último, para poder hablar de las peores consecuencias). Sin embargo, hay algo curioso en la idea de que podría haber consecuencias no comparables en una situación decisoria. Pues, aunque pueda no tener una respuesta de momento el agente, con respecto a cual de dos consecuencias prefiere, sin embargo una situación decisoria se caracteriza entre otras cosas porque *obliga* al agente a elegir una u otra consecuencia posible. Y si realmente no quiere decidirse el agente por una de dos consecuencias, la razón podría ser en el fondo que le resultan indiferentes. El otro caso posible es que el agente realmente ignora cual de las dos desea más, pero en este caso de todas formas tiene que decidir de una u otra forma cual va a considerar como más preferible (así sea mediante un volado). Por virtud de estas consideraciones, la mínima estructura que es dable exigir a la relación de preferencia es la de un orden débil. No parece haber ninguna pérdida de generalidad en suponer la existencia de (al menos) una opción máximamente preferida $c^*$ y (al menos) una mínimamente preferida $c_*$. Como resultado de esta discusión podemos introducir ahora la definición de estructura de preferencia.

DEFINICIÓN 1: Una *estructura de preferencia* es un cuádruplo $\langle C, \succsim, c_*, c^* \rangle$ que satisface los siguientes axiomas para todo $c \in C$:

(1) $\succsim$ es una relación binaria conectada, reflexiva y transitiva sobre $C$

(2) $c_*, c^* \in C$

(3) $c^* \succ c_*$

(4) $c^* \succsim c \succsim c_*$

3. *Estructura de verosimilitud*

Habíamos dicho que el agente puede o no saber, para la decisión arbitraria $d$ si la conexión entre $d$ y su consecuencia es determinista, o incluso puede ignorar si hay o no hay tal conexión. Para tener en cuenta este problema de ignorancia, habíamos sugerido introducir una idea muy psicológica de verosimilitud: una escala cualitativa, que puede ser interpretada como grado de creencia, indicando la intensidad de adhesión del agente a la creencia de que una determinada decisión $d$ acarreará una determinada consecuencia $c$. Watkins (1970) ha sugerido una forma de construir esta escala, siguiendo el ejemplo de las escalas sensoriales de calor.[1] Una escala como ésta, sin embargo, aún cuando fuese una escala numérica, no necesariamente tendría que satisfacer los axiomas usuales de la teoría matemática de las probabilidades.

El concepto cualitativo de verosimilitud recién introducido es quizá demasiado sicologista. No es necesario descartar la posibilidad de que en algunas ocasiones la verosimilitud tenga un sentido más objetivo. La pregunta interesante es más bien esta: ¿cuáles son las condiciones mínimas de estructura que debe poseer la relación de verosimilitud? Se trata de encontrar condiciones que puedan interpretarse de modo subjetivo (bayesianamente) u objetivo.

La primera cuestión que surge en conexión con el problema de determinar la estructura de la verosimilitud cualitativa es el relativo al del *espacio* en el que habrá de ser definida dicha

[1] *Cf.* pp. 195 y 196.

relación. Es obvio que habrá de ser un espacio de eventos; la cuestión estriba más bien en cómo representar este espacio. Es usual hacerlo en términos de un álgebra de conjuntos $\Im$ sobre un conjunto $X$: los eventos son los elementos de $\Im$. No parece haber pérdida importante de generalidad si se adopta esta forma de representación. Por ende, nuestro problema se reduce a las condiciones que debe de satisfacer la relación $\trianglerighteq$ de verosimilitud (o probabilidad cualitativa). Condiciones más o menos obvias son que $\trianglerighteq$ debe ser un orden débil sobre $\Im$, así como que el evento cierto $X$ debe ser estrictamente más verosímil que el evento imposible $\emptyset$. Una condición más sustancial e interesante afirma que si dos eventos excluyentes $A_1, A_2$, son respectivamente más verosímiles que dos eventos también mutuamente excluyentes $B_1, B_2$, entonces es más verosímil que ocurra uno de los dos primeros a que ocurra uno de los dos segundos. Se considera que estas condiciones (o algunas variantes de las mismas) definen lo que en la literatura sobre el tema se llama 'estructuras de probabilidad cualitativa' o —como preferimos llamarlas aquí— estructuras de verosimilitud.[2] Esta discusión puede ser resumida en la siguiente definición. Más adelante escribiremos '$A \triangleright B$' como abreviatura de '$A \trianglerighteq B \wedge \neg B \trianglerighteq A$', y '$A \bowtie B$' como abreviatura de '$A \trianglerighteq B \wedge B \trianglerighteq A$'.

DEFINICIÓN 2: Un *espacio de verosimilitud* es una terna $\langle X, \Im, \trianglerighteq \rangle$ que satisface los siguientes axiomas para todo $A, A_1$, $A_2, B_1$ y $B_2$ en $\Im$:

(1) $X$ es un conjunto no vacío, $\Im$ es un álgebra de conjuntos sobre $X$ y $\trianglerighteq$ es una relación binaria sobre $\Im$

(2) $\langle \Im, \trianglerighteq \rangle$ es un orden débil

(3) $X \triangleright \emptyset$ y $A \trianglerighteq \emptyset$

[2] *Cf.* Krantz, *et al.* (1971).

(4) Supóngase que $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$. Entonces, para cada $i$ ($i = 1, 2$)

$$A_i \trianglerighteq B_i \text{ syss } A_1 \cup A_2 \trianglerighteq B_1 \cup B_2.$$

## 4. *Hacia la teoría general*

Hemos así introducido casi todos los ingredientes conceptuales que nos permitirán formular nuestra teoría general de las decisiones. A grandes rasgos, lo que esta teoría pretende es caracterizar las situaciones decisorias de manera plenamente general, con el objeto de permitir la formulación de condiciones de racionalidad para la toma de decisiones. La tarea conceptualmente interesante consiste en encontrar condiciones plausibles (axiomas) que conecten la preferencia entre consecuencias con la verosimilitud de los eventos y la preferencia entre decisiones; esto es, lo interesante es encontrar condiciones interesantes que caractericen las opciones preferenciales en términos de la verosimilitud de los eventos posibles y las opciones preferenciales. Como se subrayó más arriba, esta caracterización no tiene que se necesariamente la utilidad esperada, sino la más general posible.

Más precisamente, uno de los conceptos fundamentales primitivos de la teoría general de las decisiones es el de preferencia entre opciones (ya mencionado arriba). El problema consiste en definir implícitamente (esto es, mediante axiomas) ėste concepto, en términos de los otros conceptos fundamentales primitivos. Esto monta tanto como determinar las leyes fundamentales de la teoría. Para tal efecto, es necesario introducir nuevos elementos conceptuales.

La consecuencia $c$ de la decisión que tome el agente depende de qué evento tenga lugar entre los elementos de conjunto fijo y exhaustivo de eventos $\mathfrak{I}$. Por ejemplo, si toma la decisión $d$ y ocurre el evento $A_1$, entonces confronta la consecuencia $c_1$; pero si ocurre el evento $A_2$ entonces confronta la consecuencia $c_2$. Lo que esto significa es que una vez tomada la decisión $d$, la ocurrencia del evento $A_1$ trae consigo aparejada necesariamente la consecuencia $c_1$, mientras que la ocurrencia de $A_2$ acarrea necesariamente la consecuencia $c_2$. Esto sugiere que a cada evento $A \in \Im$ y decisión $d$ le corresponde precisamente una consecuencia: la decisión $d$ y el evento $A$ unívocamente determinan la consecuencia $c$. Esto define una función $\kappa$ del producto $D \times \Im$ en $C$; subrayaremos en ocasiones esta conexión escribiendo $c = \kappa(d,A)$.

Dado que la elección habrá de hacerse en primera instancia entre opciones, es necesario encontrar una forma de determinar cuáles son las mejores. Los términos de este problema vienen ya acotados por el hecho de que cada decisión da lugar a una familia de posibles consecuencias, dependiendo cuál ocurre realmente del evento que tenga lugar. Si cada decisión acarreara necesariamente una cierta consecuencia (*i.e.* si la función $\kappa$ dependiese solamente de la variable $d$ y no de la variable $A$) entonces las mejores decisiones serían aquellas que arrojasen las consecuencias más preferibles (recuérdese que el conjunto de las consecuencias es acotado en lo que a preferencias concierne). Es la existencia de eventos inciertos lo que introduce las dificultades, como el ejemplo anterior de los precios lo ha mostrado. Para tratar este problema, es necesario tener en cuenta que cada decisión $d \in D$ induce una relación de verosimilitud sobre el conjunto $C$ de las consecuencias. Las consideraciones conducentes a establecer una conexión lógica entre las decisiones y las consecuencias deben tomar en cuenta no sólo las preferencias entre consecuencias, sino también las relaciones de verosimilitud inducidas por las opciones sobre ellas mismas. Cada decisión determina en efecto un espacio de verosimilitud sobre las consecuencias. Este espacio es determinado como sigue. Escribiendo la restricción de $\kappa$ a $d \in D$ como $\kappa_d$ sea $X' = \operatorname{ran} \kappa_d$; sea $\Im'$ la familia que contiene todos los conjuntos $A' \subseteq C$ tales que $\kappa_d^{-1}(A') \in \Im$; y sea $\trianglerighteq_d$ la relación definida por la condición

$$A' \trianglerighteq_d B' \text{ syss } \overset{-1}{\kappa_d}(A') \trianglerighteq \overset{-1}{\kappa_d}(B')$$

Entonces tenemos el siguiente teorema, cuya demostración se deja a cargo del lector.

TEOREMA: *Para cada decisión $d \in D$, la estructura $\langle X', \Im', \trianglerighteq_d \rangle$ es un espacio de verosimilitud.*

Por virtud de este teorema, podemos afirmar que cada decisión $d$ tiene asociada una relación de verosimilitud $\trianglerighteq_d$ sobre $C$. El problema de las conexiones lógicas entre las opciones y las consecuencias requiere alguna forma de comparación entre los espacios de verosimilitud inducidos por las opciones. Para ilustrar esta afirmación, volvamos al ejemplo de los precios. En nuestros términos, tenemos $D = \{\alpha, \beta\}$; $X = A \cup B$; $\Im = \{\emptyset, A, B, X\}$; $X \trianglerighteq B \trianglerighteq A \trianglerighteq \emptyset$; $C = \{-50, 50, 100, 150\}$; $150 \succsim 100 \succsim 50 \succsim -50$; $\kappa : D \times \Im \rightarrow C$, con $\kappa(\alpha, A) = 150$, $\kappa(\alpha, B) = -50$, $\kappa(\beta, A) = 100$ y $\kappa(\beta, B) = 50$.

En la estructura inducida por la decisión $\alpha$ en el conjunto de las opciones tenemos las relaciones

$$-50 \trianglerighteq_\alpha 150 \vartriangleright_\alpha \emptyset \bowtie_\alpha 100 \bowtie_\alpha 50;$$

en la estructura inducida por la decisión $\beta$ tenemos

$$50 \trianglerighteq_\beta 100 \vartriangleright_\beta \emptyset \bowtie_\beta 150 \bowtie_\beta -50.$$

Para poder decidir entre las opciones $\alpha$ y $\beta$, *haciendo uso exclusivo de las relaciones entre los espacios inducidos*, necesitamos comparar de alguna forma los mismos. La pregunta es: ¿cuál opción es preferible? Generalmente esta pregunta se interpreta como una petición de explicitar *principios universales de racionalidad* y de aplicar estos criterios para establecer la más "racional" de las opciones. Si intentásemos, sin embargo, explicitar alguno de estos criterios en conexión con el ejemplo que nos ocupa, veríamos que nuestras estructuras no están lo suficientemente determinadas para ello. Por ejemplo, alguien podría intentar mostrar que la decisión $\beta$ es más racional que la $\alpha$ porque garantiza una ganancia de al menos 50 millones, mientras que en la otra existe el riesgo de perder la misma suma. Dejando de lado el hecho de que a alguien le puede parecer demasiado conservador ese criterio de racionalidad, la mera posibilidad de la comparación presupone más de lo que las estructuras brindan. Pues no es lo mismo una probabilidad de .99 de perder 50 millones a otra de .01. Desafortunadamente, nuestras estructuras no hacen distinciones de este tipo: nos dicen que una opción es más verosímil que otra, pero no nos dicen "qué tanto más" lo es. Se ve, pues, que para poder tener la posibilidad de comparar las opciones a través de sus consecuencias probables hace falta introducir la noción de probabilidad cuantitativa.

Suponiendo que de alguna forma hubiésemos obtenido probabilidades cuantitativas,[3] por mor del ejemplo digamos que en la estructura inducida por $\alpha$ la probabilidad de perder 50 millones es de .75, mientras que la de ganar 150 es de .25. Digamos también que en la estructura inducida por $\beta$ la probabilidad de ganar 50 millones, al igual que la de ganar 100, es de .5. ¿Es correcto decir que es más racional la opción $\beta$ que la $\alpha$? Si esto fuera así, habría que concluir que los apostadores en los juegos de azar son irracionales y no hay duda que muchos piensan que ello es así efectivamente. Parece claro, no obstante, que lo único que está expresando este uso del término 'racional' es un juicio ético acerca de la actitudes frente a las situaciones de incertidumbre: 'irracional' es, así, toda actitud que no se conforme con cierto conservadurismo frente a los riesgos.

[3] No voy a considerar aquí el problema de las condiciones que hay que agregar a los axiomas de la Definición 2 para obtener una representación de la verosimilitud mediante una medida de probabilidad. Para esto, véase Krantz, *et al.* (1971), capítulo 5; Fishburn y Roberts (1989), y Van Lier (1989).

Me pregunto si Cristóbal Colón hubiera descubierto el Nuevo Mundo si los Reyes Católicos hubiesen sido tan "racionales".

Sostengo que es ilegítimo proclamar en general que las acciones menos riesgosas son más racionales que las demás. El concepto relevante de racionalidad aquí es uno que califica como racionales sólo aquellas decisiones que son tomadas despues de un proceso de deliberación "apropiado", que compara todos los espacios inducidos por las opciones, y que —en algún sentido difícil de precisar— adopta la decisión más "conveniente" para el agente. Parece claro que al menos en parte cuál es la decisión más conveniente depende de los gustos del agente, o de su grado de aversión a los riesgos. Es por ello que la racionalidad en la toma de decisiones es aquí expresada como una relación de preferencia entre las decisiones, precisamente para subrayar que, dentro de ciertos límites, la racionalidad es una cuestión de gustos. En otras palabras, la racionalidad admite una gran libertad en cuanto a la elección de las opciones más preferibles. El problema radica en fijar los límites donde la libertad empieza a convertirse en estupidez.

La primera restricción tiene que ver con la estructura de las preferencias sobre $D$, el conjunto de las opciones. Por virtud de las consideraciones arriba expuestas en relación con las consecuencias, la mínima condición razonable es que $D$ forme una estructura de preferencias en el sentido de la Definición 1. Este es un axioma muy sustancioso, pues afirma que hay (al menos) una opción máximamente preferible y (al menos) una menos deseable. Otro principio que apenas admite discusión es el llamado de 'la cosa segura'. Este principio podría formularse, dentro del marco conceptual desarrollado hasta aquí, como sigue: dadas dos decisiones $d, d' \in D$, si $d$ y $d'$ inducen los mismos eventos sobre $C$, *i.e.* $\Im_d = \Im_{d'}$; si, además, para cada evento $E$ en $\Im_d$, la probabilidad de $E$ en $\Im_d$ es no menor que la probabilidad de $E$ en $\Im_{d'}$; y si, más aún, al menos un evento más preferible tiene más probabilidad de ocurrir en $\Im_d$

que en $\Im_{d'}$ entonces la decisión $d$ es preferible a la decisión $d'$.

Pienso que las dos condiciones anteriores —que constituyen fundamentalmente lo que quisiera denominar una estructura general de decisiones— son lo suficientemente sustanciosas como para constituir una teoría general de las decisiones, pero surge un problema de estructura lógica en conexión con la extensión de la relación de preferencia entre puntos en $C$ a la relación de preferencia entre elementos de las álgebras sobre $C$ inducidas por la relación de verosimilitud entre los eventos de $\Im$. Originalmente, hemos definido '$\succsim$' entre elementos de $C$, no de $\Im$. ¿Cómo se puede extender esta relación para cubrir también éstos últimos? Nótese, en primer lugar, que el problema no surge tanto en los casos en que $C$ es a lo sumo numerable, pues en estos casos podemos limitarnos a considerar los eventos unitarios o elementales, *i.e.* las probabilidades *de los puntos de* $C$. Pero en los casos en los que $C$ tiene la cardinalidad del continuo (que es efectivamente el otro único caso considerado por la teoría), puede suceder que los puntos tengan probabilidad nula, más no así algunos de los conjuntos de Borel que los contienen. En este caso, de la misma manera que se requiere una densidad de probabilidad sobre $C$ para, integrando, obtener una medida de probabilidades sobre sus subconjuntos de Borel, también se requiere algo así como una "densidad de utilidad" para, integrando, obtener una medida de utilidad sobre los mismos borelianos. En otras palabras, si $E \in \Im_d$ queremos definir la utilidad $u(E)$ de $E$ mediante la ecuación

$$u(E) = \int_{c \in E} u(c)dc.$$

Es inmediato que si esta integral existe y es finita para cada $E \in \Im_d$, entonces $u$ es una medida aditiva sobre $\Im_d$; pues si $E, F \in \Im_d$ con $E \cap F = \emptyset$, entonces

$$(E \cup F) = \int_{c \in E \cup F} u(c)\, dc = \int_{c \in E} u(c)\, dc + \int_{c \in F} u(c)\, dc.$$

No es éste el lugar para considerar las condiciones bajo las cuales estas integrales están bien definidas. En vez de ello, quisiera decir unas palabras acerca de la forma que asume en el marco conceptual introducido la diferencia entre la racionalidad sustantiva y la procesal.

La racionalidad sustantiva se podría caracterizar por exigir básicamente dos cosas: (1) la definición de una función objetivo sobre $D$, que represente la relación de preferencia entre los elementos de $D$; y (2) la explicitación de un algoritmo que permita maximizar la función objetivo en términos de una comparación sistemática de todas las estructuras inducidas sobre $C$. En contraste, la racionalidad procesal exige una "deliberación apropiada" que permita comparar, quizá de manera incompleta y asistemática, las diferentes estructuras inducidas, con la esperanza no de obtener la maximización de una función objetivo como la ya descrita, sino más bien de encontrar "buenas" opciones. En este respecto, es difícil proporcionar reglas más generales que las restricciones ya anotadas (que la relación de preferencia entre las opciones cumpla las condiciones de la Definición 1, y que satisfaga el principio de la cosa segura), pues hay una infinidad de tales relaciones, dependiendo de los gustos de los agentes, de su aversión al riesgo, de las consideraciones externas relevantes que puedan traer a colación al momento de construir su relación sobre las opciones, etcétera. El problema fundamental de la teoría general de las decisiones es más bien éste: determinar en cada caso la función de preferencia entre opciones proporcionando criterios objetivos para decidir entre estructuras inducidas, de la forma más sistemática posible. Este es un ejercicio no trivial en la mayoría de los casos y que —a mi modo de ver— constituye la parte más enjundiosa de la teoría de las decisiones.

Para concluir, quisiera indicar que la teoría de las decisiones no es una teoría "explicativa", sino más bien un marco conceptual dentro del cual es posible plantearse problemas de decisión. Sin embargo, puede utilizarse para proveer explicaciones acerca de la conducta de un agente si se conoce su relación de preferencia, con respecto a ciertos eventos y consecuencias, a guisa de un hábito del mismo agente. Es decir, si conocemos las preferencias de un agente como un hábito del mismo, entonces podemos predecir (y por ende explicar) qué decisión habría de tomar bajo ciertas circunstancias. Aquí el término inicial de la explicación es un hábito, es decir, un accidente inherente en la persona del agente, cuya existencia, a la vez, podría ser explicada en términos de su historia personal pasada.

## REFERENCIAS


Borger, R. y Cioffi, F., *Explanation in the Behavioural Sciences*, Cambridge UP, Cambridge, 1970.

Fishburn, P. C., y Roberts, F. S., "Axioms for Unique Subjective Probability on Finite Sets", en *Journal of Mathematical Psychology 33*, pp. 117–130 (1989).

Hahn, F. y Hollis, M., *Filosofía y teoría económica*, FCE, México, 1986.

Krantz, *et al*, *Foundations of Measurement I*, Academic Press, New York, 1971.

Simon, H. A., "De la racionalidad sustantiva a la procesal", en Hahn y Hollis (1986).

Van Lier, L., "A Simple Sufficient Condition for the Unique Representability of a Finite Qualitative Probability by a Probability Measure", en *Journal of Mathematical Psychology 33*, pp. 91–98, (1989).

Watkins, J., "Imperfect Rationality", en Borger y Cioffi (1970).


*Recibido: 11 de septiembre de 1990*

## SUMMARY

Starting from Simon's (1976) opposition between procedural and substantive rationality, the paper addresses the question whether it is possible to characterize the former in general terms, and discusses the problem of overcoming the aforementioned opposition. It tries to show, by means of the concept of a general decision structure, that even though it is not possible to characterize procedural rationality in an exhaustive and fully general way, at least it is possible to introduce a scheme within which it is possible to pose problems of procedural rationality, or even to postulate some axiomatic restrictions for the same. It is not hard to see that substantive rationality is a particular case of this scheme, namely, the case in which it is possible to *measure* certain theoretical parameters and adopt as rules of decision certain classical principles grounded upon the concept of expected utility. The paper does not provide, however, the details that show how the special theory is obtained out of the more general one.

*[A. G. S.]*